Behavior Profiling and Analysis in Wireless Home Networks

Kuai Xu, Feng Wang
Division of Mathematical and Natural Sciences
New College of Interdisciplinary Arts & Sciences
Arizona State University

Bin Wang
Department of Computer Science and Technology
College of Information Science and Engineering
Shandong Agricultural University

I. Introduction

In recent years, the low cost of wireless technologies and residential broadband networks have driven the wide deployment of wireless home networks (WHNs). The ubiquitous availability of WHNs enables users the access to the Internet from everywhere within their homes. However, it also opens the doors for the drive-by hackers that exploit open access home networks for Internet connections [1]. Previous work such as [2] and our own measurement studies have shown the existence of a large amount of open or un-encrypted access points in wireless residential networks. For example, our recent measurement experiment finds an average of 35% are open home wireless networks in six residential buildings. At the same time, Internet attackers actively explore vulnerable home computers and turn them into part of botnets for sending spams or launching distributed denial of service (DoS) attacks [3].

The existing wireless access pointers from commercial vendors such as Linksys and Netgear are mostly built with NAT solutions and stateful packet inspection firewalls [4]. These techniques are very useful to filter attacks with known patterns, however they lack the ability to detect novel attacks or existing attacks with new variations. Therefore, it is very important to develop behavior-oriented techniques that do not rely on signatures for detecting such attacks.

In this short paper, we present a preliminary design of a behavior profiling system in WHNs for network security monitoring. Figure 1 illustrates a schematic architecture of the behavior profiling system that is deployed in a typical wireless home network. The goals of the proposed behavior profiling system are to i) actively learn the traffic patterns of wireless home networks, ii) detect anomalous behavior from inside networks as well as from the Internet. Based on network traffic patterns for each computer, the system builds baseline behavior profiles, and subsequently detects events of interest through behavior deviations. The contributions of this work are two-fold. First, we propose to build the behavior profiles for each computer in WHNs towards a deep understanding of the traffic patterns in wireless residential networks. Secondly, we present a systematic architecture that aims to detect anomalous behavior through real-time traffic profiling.

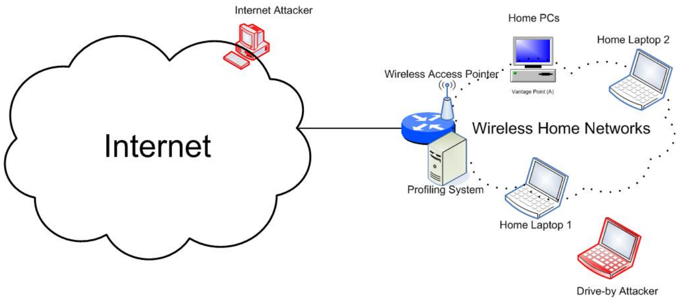

Fig. 1. Behavior profiling system for wireless home networks.

The reminder of this short paper is organized as follows. Section II presents our behavior profiling methodology that builds a baseline profile for each computer within the home network as well as the aggregated network itself. Section III briefly describes the initial design of the profiling system. Finally, Section IV summarizes this paper and outlines our ongoing and future work.

II. Behavior Profiling Methodology

A. Profiling Wireless Home Networks

The basic building block for understanding the behavior of a home computer in WHNs is its communication pattern, such as *what applications does it use*, *what is the typical traffic load*. To quantitatively study such patterns, we use *entropy* from information theory to measure the *distributions* of the feature dimensions, e.g., IP addresses, port numbers, and protocols, from all the applications of a home computer. Our focus on the distributions, rather than the volume of the dimensions is motivated by our previous study and other work, which show that the distribution measures are more efficient for detecting anomaly traffic behavior in IP networks [5], [6].

To analyze the behavior of each legitimate computer, we first separate its inbound traffic based on the applications, and measure the distributions of the traffic features. For example, for the source IP addresses of a given application x, we could compute the entropy as $\mathcal{H}(x) = \sum_{i=1}^{m} p_i \log p_i$, where m denotes the total number of source IP addresses that sends traffic of the application x to the computer, and p_i denotes the percentage of packets from the i-th IP address. Subsequently, we use the weighted sum to calculate the average entropy of the source IP addresses across all applications:

978-1-4244-5176-0/10/$26.00 ©2010 IEEE

$\mathcal{H} = \sum_{x=1} \mathcal{N} \frac{\mathcal{V}_x}{\mathcal{V}_{total}} \mathcal{H}(x)$, where $\mathcal{N}$ denotes the total numbers of applications and $\mathcal{H}$ represents the average entropy of the source IP addresses for the computer. Similarly, we could compute the entropy of the other feature dimensions for the computer based on its communication patterns, such as source port, destination ports, and packets, etc. These entropy distributions carry a rich information on the communication patterns of these home computers and form their behavior profiles.

The profile of the entire WHNs maintains the aggregated information of individual home computers. In our study, we are interested in collecting the following metrics: the number of home computers, inbound and outbound packets, and bandwidth utilizations. These information provides an aggregated view of the network, and is of great value to detect the drive-by attackers, since the computer used by the attacker likely changes the statistics of host counts and bandwidth utilizations.

B. Detecting Anomaly Behavior

The behavior profiles of home computers and WHNs provide an informative baseline for detecting anomaly behavior in WHNs. To detect the attacks traffic towards wireless home computers, a simple and effective method is to calculate an weighted exponential moving average of the entropy, $\mathcal{H}'$ as the baseline behavior: $\mathcal{H}'_i = \alpha * \mathcal{H}_{i-1} + (1-\alpha) * \mathcal{H}'_{i-1}$. A large deviation, ϵ, of the observed entropy, $\mathcal{H}$, from the baseline, $\mathcal{H}'$, $\epsilon = |\mathcal{H} - \mathcal{H}'|$, could indicate an interesting network event.

As the profile of WHNs includes the total number of inside computers, the appearance of drive-by attackers could immediately trigger an alarm, since any traffic from the drive-by attackers have to also traverse through the wireless access point and are captured by the behavior profiling system. On the other hand, the drive-by attacker could employ spoofing attacks by sending traffic with one of the home computers to the Internet. In this scenario, the profile of the spoofed computer could help detect the attack behavior, since the returning traffic that responds to the spoofing traffic changes its normal behaviors.

III. System Architecture

To evaluate the operational feasibility in real WHNs environment, we design a behavior profiling system architecture which consists of four key components: application configurations, data collection, behavior profiling, and anomaly reporting. Below we briefly describe the major functions of each component.

The application configuration component allows the users of WHNs to input the known services running in the networks. Such user-generated information are very valuable for validating the effectiveness of the system.

The major function of the traffic collection component is to collect IP data traffic that traverse through the wireless access point. Our current implementation is to install Wireshark on a desktop that connects to the same local area networks with the access point and listen to the traffic through the shared medium, thus any traffic between WHNs and the Internet will be captured by the collection component.

Behavior profiling is the key component of this system, and it implements the methodology described in the previous section. This component establishes the baseline of normal traffic patterns for each computer in WHNs through active learning. The baseline is further employed to detect significant deviations that are caused by anomalous behavior towards these computers. In addition to the profiles of individual computers, the behavior profiling component also maintains a summary of the network during each time period, e.g., number of computers, number of packets, number of bytes. The last component, anomaly reporting, is to report the detected anomaly behavior to the users.

IV. Conclusions and Ongoing Work

The prevalence of wireless technologies and residential broadband networks bring accessibility and convenience to millions of home networks. However, the ubiquitous availability of wireless home networks also brings opportunity for the drive-by attackers who could connect to these access networks through open wireless connections without physically being inside the homes. In addition, many Internet attackers continuously scan vulnerable home computers and control them as part of botnets for sending spams or launching denial of service attacks. In this short paper, we presented a preliminary design of a behavior profiling system that attempts to build behavior profiles of wireless home networks for security monitoring. We are currently in the process of implementing a real-time behavior profiling system. Our future work focuses on i) the evaluations of detection accuracy, e.g., false positive and false negative through trace-driven simulations, and ii) the deployment of this system in a number of home networks to evaluate the effectiveness of this system in detecting attacks from the home network as well as from the Internet. In addition to real Internet traffic, we plan to simulate a variety of attacks to evaluate the detection accuracy of this system.

References

[1] Alex Tsow; Markus Jakobsson; Liu Yang; Susanne Wetzel, "Warkitting: The Drive-by Subversion of Wireless Home Routers," *Journal of Digital Forensic Practice*, vol. 1, p. 179 192, 2006.

[2] D. Han, A. Agarwala, D. Andersen, M. Kaminsky, K. Papagiannaki, and S. Seshan, "Mark-and-Sweep: Getting the Inside Scoop on Neighborhood Networks," in *Proc. of ACM Internet Measurement Conference*, October 2008.

[3] Z. Li, A. Goyal, Y. Chen, and V. Paxson, "Automating analysis of large-scale botnet probing events," in *Proc. of International Symposium on Information, Computer, and Communications Security*, March 2009.

[4] Cisco Linksys, "Protecting Your Network," http://www.linksysbycisco.com/US/en/learningcenter/ProtectingYourNetwork.

[5] T. Karagiannis, K. Papagiannaki, N. Taft, and M. Faloutsos, "Profiling the End Host," in *Proc. of Passive and Active Measurement Conference*, April 2007.

[6] K. Xu, Z.-L. Zhang, and S. Bhattacharyya, "Profiling Internet Backbone Traffic: Behavior Models and Applications," in *Proceedings of ACM SIGCOMM*, August 2005.